Platform Intelligence in News (PIN)

a research and development project on artificial intelligence in the news industry

Insights and recommendations for news publishers

AUTHORS

Kasper Lindskow
Head of AI at JP/Politikens Media Group and
project manager for Platform Intelligence in News

Mikkel Flyverbom
Professor at Copenhagen Business School,
dept. of Management Society and Communication)

Árni Einarsson
Tenure track assistant professor at University of Copenhagen,
dept. of Communication

Stine Lomborg
Professor at University of Copenhagen, dept. of Communication

Rasmus Helles
Associate professor at University of Copenhagen,
dept. of Communication

Nanna Thylstrup
Associate professor at University of Copenhagen,
dept. of Communication

Anders Søgaard
Professor at University of Copenhagen,
dept. of Computer Science

Jes Frellsen
Associate professor at the Technical University of Denmark,
dept. of Applied Mathematics and Computer Science

November 2024

CONTENTS

1. INTRODUCTION 4

2. WHY THE PIN PROJECT? 5

3. ALIGNING AI WITH NEWS PUBLISHING – PIN'S VALUES COMPASS 6

4. HOW CAN NEWS PUBLISHERS USE AI? 7

5. KEY FINDINGS AND INSIGHTS 8

6. RECOMMENDATIONS TO NEWS PUBLISHERS 16

7. QUESTIONS ON THE HORIZON 21

8. WANT TO LEARN MORE ABOUT AI IN MEDIA? 22

1. INTRODUCTION

This report presents insights and recommendations to news publishers on how to integrate artificial intelligence (AI) in their work processes and news experience from the Platform Intelligence in News project (PIN). PIN was a comprehensive applied research project that aimed to develop effective and responsible AI systems for news. The project was a collaboration between JP/Politikens Media Group and three universities that began in October 2020 and concluded in June 2024 and received funding from Innovation Fund Denmark.

AI is poised to impact the news industry in years to come, and this development holds both a wealth of potentials and a wealth of pitfalls in need of consideration. The promises of AI solutions for the news industry are about optimizing and enriching news production, distribution, recommendation and monetization. The challenges and dangers for the news industry involve increased dependency on commercial tech platforms and loss of control, the undermining of business models and problems with bias and credibility of information.

News publishers face a wealth of strategic choices and decisions in the coming years. The most important are concerned with technology, journalistic mission and values, business models, and their relationship with their users. Hence, the findings and recommendations in this report are categorized along these four domains. The central argument in this report is that the news industry must strive for alignment between its core tasks and AI tools, and this can be achieved by engaging with the 15 insights and 15 recommendations presented here. The figure below summarizes the report's insights and recommendations which are unfolded in sections 5 and 6.

Figure. *Overview of insights and recommendations.*

		Insights (section 5)	Recommendations (section 6)
Technology	1	AI technology is diverse and fast evolving	Monitor developments in AI technology continuously
	2	AI systems must be adapted to the news domain	Don't just copy AI solutions from companies in other sectors
	3	New types of quality controls are needed	Introduce domain specific quality controls
Journalism	4	AI technology is not neutral	Monitor the effects of AI systems on your news content and news flow
	5	Alignment with editorial profile requires constraining AI systems	Align your AI systems with your editorial values
	6	Integrating editorial values into AI systems is an ongoing challenge	Always have editors in the loop – albeit at different levels of abstraction
Business	7	AI can increase the revenues of news publishers	Invest in mature AI cases to increase revenues – but also allow experimentation if you strive for competitive advantage
	8	Generative AI can produce tools that assist with tasks in the editorial workflow	Invest in AI to increase productivity – but do not expect large returns in the short term
	9	AI leads to dependencies on tech giants in some AI domains and not in others	Manage your dependencies
Users	10	Users interpret recommender systems via their experiences on social media and expectations to the publisher	Understand what people value
	11	The (perceived) introduction of personalization changes the relationships between news publishers and readers	Recognize that users' behaviors and experiences are complementary
	12	Recommender systems have impact on exposure, but effects are small	Manage how AI systems change relationships with publishers
Organization	13	Interdisciplinary collaboration is necessary for developing healthy AI systems	Take a strategic and adequately funded approach
	14	Knowledge brokers are essential for the successful adoption of AI	Nourish interdisciplinary collaboration when designing AI systems
	15	Cross-industry knowledge sharing fosters responsible innovation	Engage in collaborate with peers across the news industry

2. WHY THE PIN PROJECT?

News publishers have undergone massive digitalization over the past three decades. This development has come in waves: from analog to digital, from web 1.0 to web 2.0, from desktop to mobile, etc. Artificial intelligence seems poised to become the next major wave of digital transformation to hit the media industry.

In the past waves of digitization, most news legacy publishers were reactive takers of technology, rather than proactive drivers of transformation. While this did enable innovation to the benefit of publishers and news readers, most parts of the media industry became heavily dependent on digital solutions developed by big tech companies with their own commercial goals. In the process, publishers opened their businesses to data harvesting and allowed social media platforms to distribute and monetize their news content. As a result, many parts of the news business have been overrun by tech giants that harvest and commercialize data, dominate advertising, and largely set the terms and conditions of digital spaces in ways that challenge the business models that journalism have traditionally rested on.

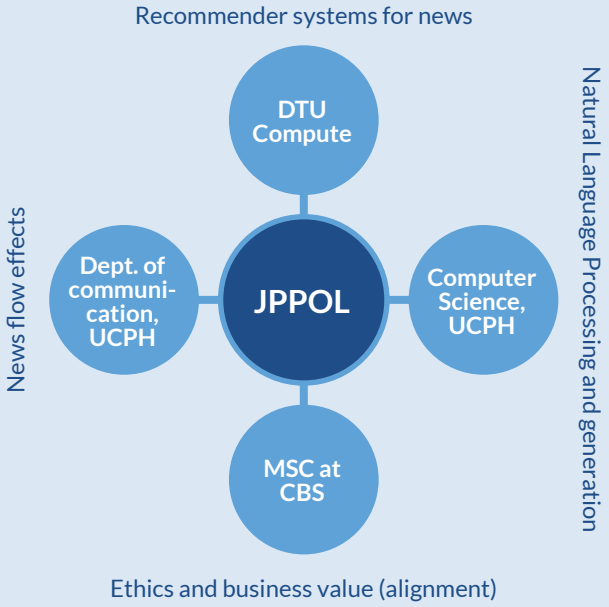
The current wave of AI-driven systems and technologies have the potential to impact journalism even deeper than earlier waves, as they potentially affect all stages of the news publishing value chain, including news discovery, production, distribution, and monetization. While the exact scope, timing, and ramifications are still uncertain, it is likely that a reactive position towards AI would favor large technology companies over news publishers and journalistic values.

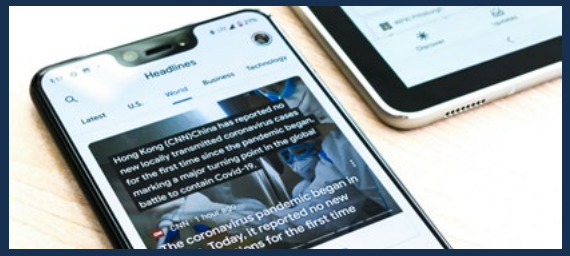
AI has already proven that it can be used to enhance user experiences and value creation in digital media, including video, audio streaming, and social media, via technologies such as metadata production, recommender systems, and generative AI. However, the introduction of AI has also unleashed serious new problems related to bias, diversity, accountability, and disinformation, as well as broader concerns about the effects for problems related to the lack of protection, harmful contents and addictive design.

In a news publishing context, AI is argued to offer both great opportunities and significant challenges to news publishers: AI could strengthen journalism via AI powered investigations across piles of documents, seamless versioning of news stories across multiple modalities, and personalization that allows important niche stories to find their audience. It could bring economic relief to news publishers by increasing reader engagement and revenues and raising productivity in the newsroom via new editorial tools. Conversely, AI done wrong could lead to bias and factual errors that erode the credibility of news publishers, result in loss of editorial control with the news flow and increase audience fragmentation in time where many societies are already polarized. Further, it could create new dependencies of news publishers on the tech giants that accelerate the erosion of news publishers' already hard-pressed business models.

Against this backdrop of potentials and challenges, the purpose of the PIN-project was to develop responsible and effective AI systems that create value for news publishers in ways that align with their values. Further, PIN aimed to contribute to healthy norms for AI in media in the process.

The ambition of the project was to address a concern raised by many organizations and companies, namely how they get value out of AI solutions without undermining their own foundations and jeopardizing the core tasks and values that they are committed to as well as their legitimacy and position in society.




Headlines
Latest
World
Business
Technology
Hong Kong (CNN)China has reported no new locally transmitted coronavirus cases for the first time since the pandemic began, marking a major turning point in the global battle to contain Covid-19.
Today, it reported no new
for the first time

WHAT WAS THE PLATFORM INTELLIGENCE IN NEWS (PIN) PROJECT?

Platform Intelligence in News (PIN) was an applied research project that aimed to develop responsible and effective AI systems for news media that are independent of the tech giants. The project received support from the Innovation Fund and was a partnership between JP/Politikens Media Group and four departments at three Danish universities, including Department of Management, Society and Communication at the Copenhagen Business School, Computer Science at the University of Copenhagen, DTU Compute at the Technical University of Denmark, and Department of Communication at the University of Copenhagen.

Starting in October 2020, the project set out to develop strategic directions for AI, concrete technological solutions, as well as carrying out empirical research and live tests of these solutions to understand the impact of artificial intelligence on the workings of news publishers. The purpose of the PIN project was to help publishers navigate the AI wave successfully by addressing the technical, economic, journalistic, and user-related opportunities and challenges associated with the introduction of AI in the media. The project started in October 2020 and concluded in June 2024.

3. ALIGNING AI WITH NEWS PUBLISHING – PIN'S VALUES COMPASS

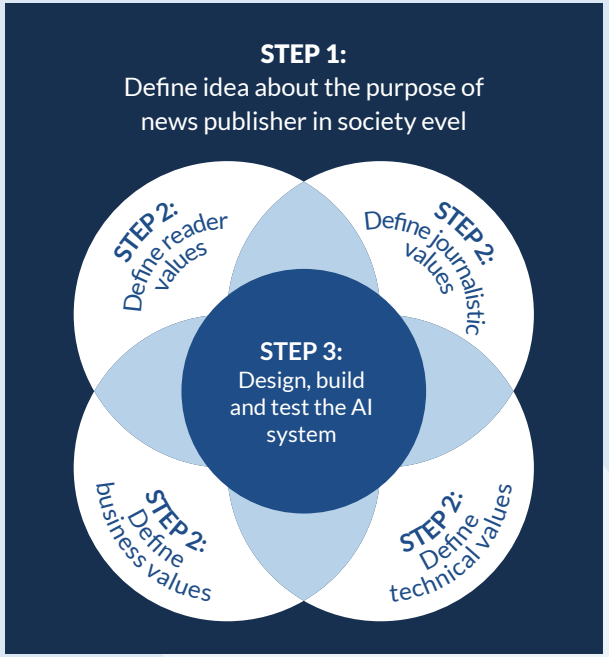
To realize the potentials and address the challenges associated with AI, PIN's point of departure was that AI systems for news should both contribute to commercial value creation and be aligned with the situated values of news publishers in democratic societies.

The notion of alignment with the values of news publishers is central to responsible use of AI. Accordingly, PIN's guiding principle was that AI systems for news publishing should be in line with general ethics principles for AI, such as transparency, fairness, preservation of the right to self-determination, and the prevention of harm. However, we found that to make these abstract principles meaningful and actionable they had to be contextualized for news publishing. To do so, we developed a values compass both as a framework for PIN and to provide guidance to practitioners in the news industry.

The value compass is both an invitation to launch a simple process that can support the design and evaluation of AI systems for news, and an invitation to engage with a specific set of concerns that must be balanced to design responsible AI systems in the news domain.

The process involves three steps, illustrated in the figure to the right:

1. Define and concretize your purpose and mission as a news publisher in society
2. Articulate how AI can help realize your purpose (1) and define the legitimate and sometimes diverging concerns that need to be considered and balanced when designing the relevant AI systems, namely your journalistic values, user-related values, business values, and technical values
3. Make sure that the decisions you make are reflected in the AI systems or evaluation frameworks that you build or buy.



4. HOW CAN NEWS PUBLISHERS USE AI?

AI is a diverse and fast evolving field that spans numerous types of applications based on machine learning. PIN chose to focus on AI technologies that were assessed to have most potential for legacy news publishers with a core focus on text-based news publishing and to become industrially mature within a horizon on two to three years. They included:

1. **Recommender systems focusing on personalization of news feeds**
   PIN produced recommender systems that were optimized for news publishing. These methods were live tested on www.ekstrabladet.dk to monitor their effects on hard metrics such as clicks, time spent reading, and subscriptions sales as well as for effects in the news flow in order to support editorial alignment. Further, PIN carried out qualitative studies of how users perceive and react to the introduction of recommender systems in the news flow.

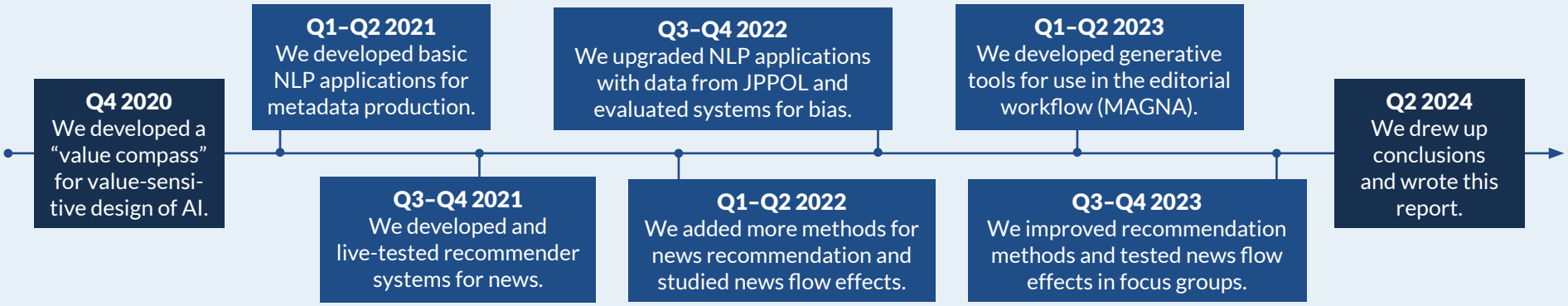
2. **AI for metadata production for news content:**
   PIN produced and live-tested a range of AI systems that generate metadata about news articles, including identifying news topics, sentiments, entities and assessing the lifespan of news articles. Further, PIN produced AI systems that create news article representations (embeddings) which proved useful for non-personalized news recommendations and generative AI systems that synthesize facts from earlier news coverage in new articles or story elements.

3. **Generative AI (chatbots) for use in the editorial workflow**
   PIN developed and live tested editorial generative AI systems based on chatbots that automate tasks to assist journalists in their production process. This included editorial tools that enable journalists to 1) use generative AI when producing news articles 2) use generative AI when covering live events, and 3) use generative AI to produce novel news articles or story elements (e.g. fact boxes or story summaries based on multiple news articles) drawing on content in the news archive via a so-called retrieval augmented generation (RAG) solution.

METADATA PRODUCTION FOR NEWS IN TEXT, VIDEO, OR AUDIO?

PINs development and testing of AI systems for news took the needs of JP/Politikens Media Group as its point of departure. As JP/Politikens Media Group is mainly a text-driven news publisher, PIN focused on natural language processing (NLP), which is a text-focused AI technology, as the core method for metadata production. To address the needs of news publishers with core focus on video or audio, AI technologies such as computer vision, speech to text (STT) and text to speech (TTS) would likely have been more central to power metadata production.

PROJECT TIMELINE


Q4 2020
We developed a "value compass" for value-sensitive design of AI.
Q1–Q2 2021
We developed basic NLP applications for metadata production.
Q3–Q4 2021
We developed and live-tested recommender systems for news.
Q3–Q4 2022
We upgraded NLP applications with data from JPPOL and evaluated systems for bias.
Q1–Q2 2022
We added more methods for news recommendation and studied news flow effects.
Q1–Q2 2023
We developed generative tools for use in the editorial workflow (MAGNA).
Q3–Q4 2023
We improved recommendation methods and tested news flow effects in focus groups.
Q2 2024
We drew up conclusions and wrote this report.

IMPLEMENTATION OF PINS AI SYSTEMS IN JP/POLITIKENS MEDIA GROUP

Most of the AI systems developed by the PIN project were subsequently implemented at scale by Ekstra Bladet, including recommender systems for news flow personalization, multiple systems for metadata production, and a suite of generative tools that assist journalists in the daily news production.

In January 2024, JP/Politikens Media Group established a shared AI unit (JPPOL AI) tasked with rolling out AI systems from PIN across the group and continuing the development of new AI systems that build on PINs learnings and exploit the new AI technologies as they emerge. The first solutions based on technology from PIN that were rolled out across the group in collaboration with each news brands' editorial rooms were certain AI systems for metadata-production and a suite of generative editorial tools called MAGNA (Monitoring and Assisted Generation of News Artefacts) that enables journalists to use generative AI in news production in ways that align with each brands specific mission and tone of voice.

5. KEY FINDINGS AND INSIGHTS

In the process of developing and live-testing AI systems in JP/Politikens Media group, the PIN project produced knowledge about the challenges and opportunities related to the use of AI in news publishing. In this section, we describe key findings related to the four domains described in section 3, as well as insights related to processes and organization that cut across the domains. The findings stem from research carried out via the PIN-project and most has been published in the academic papers that are referenced below and listed together in section 8.

TECHNICAL FINDINGS AND INSIGHTS

1. **AI technology is diverse and fast evolving**
   AI technologies, ranging from text and image processing to personalization and predictive uses, are rapidly evolving. Core examples from the past four years include the maturing of AI-based language models (transformers) which in the early 2020s enabled advancements in a certain type of recommender system (so-called content-based recommenders) which led to significant improvements in news personalization but also introduced new challenges related to potential audience fragmentation. Later in 2022, the emergence of the GPT3-generation of large language models made editorial tools based on generative AI feasible but also introduced new challenges related to e.g. content factuality which are particularly relevant in a news publishing context. This dynamic is likely to persist in the coming years, continuously blending new opportunities and challenges for publishers stemming from AI developments.

2. **AI systems must be adapted to the news domain**
   AI systems solve the tasks they are designed for with a greater or lesser accuracy. For that reason, the types of AI systems that are used to recommend movies on Netflix or generate replies to customers in chatbots will often not work well for news publishers. Rather, making AI work for news publishing requires designing systems that take domain specific challenges such as the rapid pace of the news flow and the style of writing and minimal tolerance for factual errors in news articles into account.

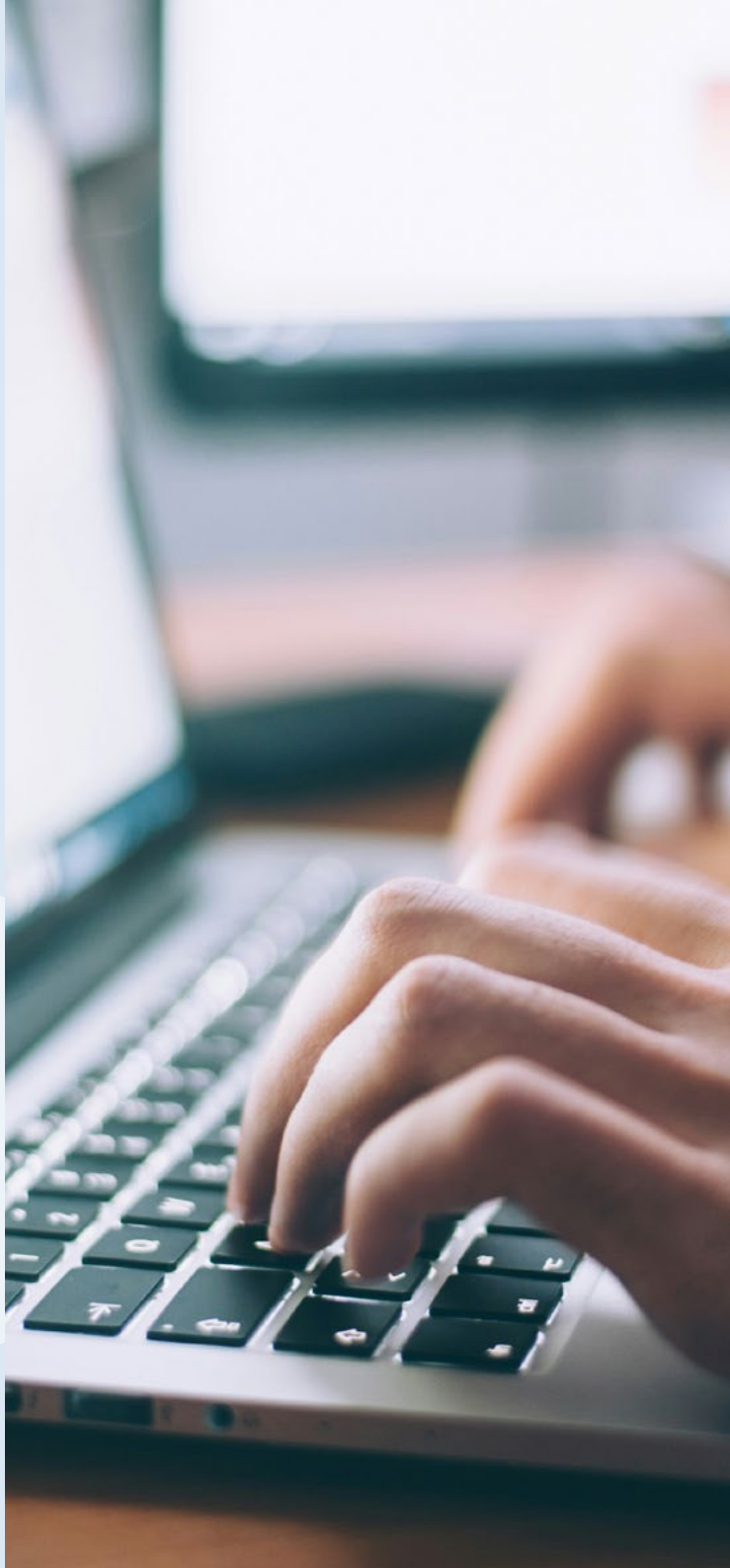
   Examples of this (encountered on the PIN-project) include generative AI systems that must be prompted in certain ways to fit the editorial profile of news publishers and implemented in certain ways (see "external grounding" below) to reduce the risks of factual errors which are particularly challenging in news coverage that strives to be factually correct. Another example was news personalization were the methods typically used in video streaming and e-commerce (so-called collaborative filtering recommenders) were found to fail in many news recommendation scenarios as they could not learn readers' interests in newly published articles fast enough. Conversely, recommender systems that personalize based on deep reader profiles and article metadata (so-called content-based recommenders) were found able to effectively recommend articles news articles instantly when they are published and proved robust in most news recommendation scenarios.

3. **New types of quality controls are needed**
While the technical systems that have traditionally enabled digital news publishing have at their foundation been logical, the machine learning that powers AI is fundamentally probabilistic. A consequence is that AI systems can perform tasks that go way beyond the capabilities of classical software systems, but also that AI systems can sometimes fail to predict the correct reader interests or 'hallucinate', producing factually incorrect information. For that reason, AI requires new types of quality control and risk mediation that must take the special needs of the news domain into account.

Avoiding factual mistakes, for example, is a key value in news publishing. As is well-known, the probabilistic nature of generative AI systems such as ChatGPT causes them to sometimes produce false information. However, challenges with hallucinations can be greatly reduced when generative AI is grounded in external facts such as news article drafts or news article archives. Nevertheless, even though such methods greatly reduce the risk of generative AI producing factual errors, editorial checks continue to be necessary to ensure correctness.

LIVE A/B TESTS OF RECOMMENDER SYSTEMS ON PIN

PIN carried out more than live 70 A/B tests of numerous recommendation methods and non-personalized content selection methods (e.g. article recommendations based on semantic search) in three phases that unfolded in 2021, 2022 and 2023. The results of these A/B tests – as well as data on the performance and effects of recommender systems rolled out at scale – are the basis for multiple insights in this report. Some of the results are more fully reported in Kruse et al 2023 while Kruse et al (2024a) and Kruse et al (2024b) reports the results of offline tests made by PIN as well as participants in ACM RecSys Challenge 2024 which was based on a news recommendation dataset from PIN and included more than 200 researchers and data scientists from around the world (read more about the challenge on recsys.eb.dk)



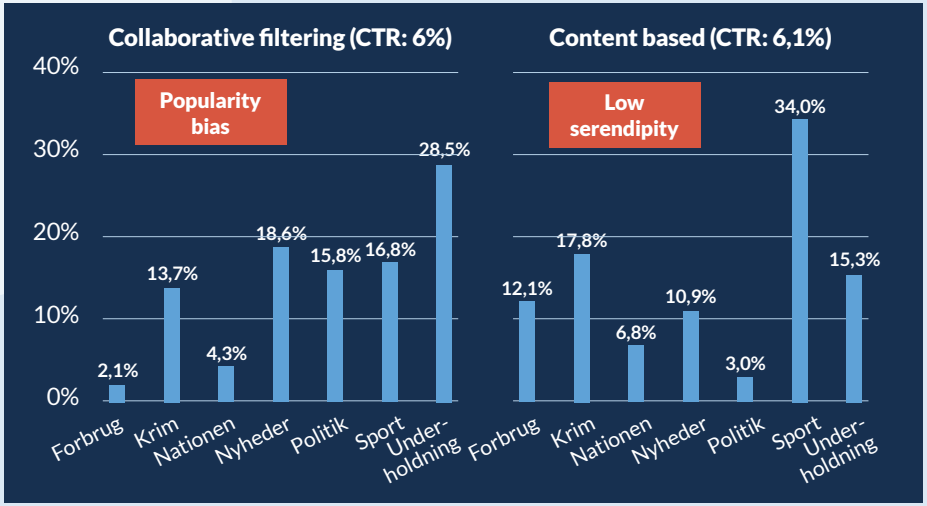
JOURNALISTIC FINDINGS AND INSIGHTS

4. AI technology is not neutral

At the most basic level, AI technologies are not "neutral machines". Rather, the choices that are made when AI systems are designed have great consequences for their performance and their effects on the production and distribution of news content.

This applies to the AI systems that produce metadata which are sensitive to among other things the taxonomies and annotation protocols chosen by the systems designers (Thylstrup and Møller-Hartley 2024). It applies to generative systems which are sensitive to both choice of model (e.g. GPT4 or Claude 3) and the prompts used when generating content, as exemplified in insight #5 below. Further, it applies to recommender systems that produce very different news flows depending on choices made about what types of recommendation methods to use and how to train them. To exemplify the latter, the figure below shows how two fundamentally different recommendation methods (collaborative filtering vs content-based) has the same ability to predict readers interests (measured by the users' click-through-rate) but produces two very different news flows (measured by news topic shares in the personalized news flows).

Figure:. *Example of the different news flow effects of two well-optimized recommender systems*



Source:. *A/B test of recommendation methods on ekstrabladet.dk from the PIN-project.*

5. Alignment with editorial profile requires constraining AI systems

Aligning AI systems with editorial values require careful thought and often necessitates implementation of constraints that regulate the AI systems' behavior. For that reason, editorial alignment of AI systems is still very much a human discipline that requires connecting situated knowledge of editorial values with technical skills in data science.

Examples (encountered by PIN) are abounding:

- The design of AI systems that produce metadata or evaluate the output of generative AI systems requires taxonomies and annotation protocols with strict definitions that necessitates deep knowledge of the news publishing domain as well as Danish cultural and social structures.

- News recommender systems, when left to optimize towards reader interests only, can easily deviate from a news publishers' editorial profile which has traditionally always balanced readers interests against editorial values. On PIN, we found is was necessary to introduce a set of rules that constrained the types of news stories that a recommender system could recommend in a personalized news flow. Further, we designed the recommender systems to train on a users' aggregate behavior on ekstrabladet.dk rather than only on the users' interaction with the recommender system element to avoid a feedback loop that could increase audience fragmentation.

- Generative AI systems tend to gravitate towards an average writing style rather than the tone of voice and journalistic angles of news publishers. We found that this can to some extent be addressed (via extensive prompting or fine tuning), as exemplified in the box below. However, aligning the output of generative language models with the editorial voice of specific news brands is likely be an ongoing challenge in generative editorial tools.

THE FIRST PARAGRAPHS OF ARTICLE DRAFTS PRODUCED BY PINS MAGNA FOR TWO DIFFERENT NEWS BRANDS IN JP/POLITIKENS MEDIA GROUP (IN DANISH)

POLITIKEN

Kamala Harris' Valgkamp: En kamp for middelklassen og modstand mod Trump

Kamala Harris, Demokraternes præsidentkandidat til valget i november, har markeret sig med en række politiske løfter og en stærk kampagne, der fokuserer på at forbedre forholdene for den amerikanske middelklasse og udfordre Donald Trumps politik. Harris' valgkamp har allerede skabt betydelig opmærksomhed og debat i USA.

Økonomiske løfter og skattepolitik

Harris har fremlagt en række økonomiske planer, der sigter mod at hjælpe middelklassen og de laveste indkomstgrupper. Hun har lovet skattelettelser til familier og lavtlønnede, og hun vil øge skattefradraget for børnefamilier.

(… draft continues, 588 words in total)

EKSTRA BLADET

Kamala Harris' valgkamp: En kamp om milliarder og støtte fra stjerner

Kamala Harris' valgkampagne har samlet en halv milliard dollars

USA's vicepræsident Kamala Harris har samlet omkring en halv milliard dollars ind til sin præsidentkampagne. Det svarer til cirka 3,37 milliarder danske kroner og er på størrelse med bruttonationalproduktet i Mikronesien. Kampagnemidlerne er afgørende i det amerikanske præsidentvalg, hvor de bruges til reklamer på sociale medier og andre platforme.

Støtte fra Taylor Swift og Stevie Nicks

Kamala Harris har fået opbakning fra store kulturpersonligheder som Taylor Swift og Stevie Nicks.

(… draft continues, 348 words in total)

Source:. *PINs generative AI toolbox, MAGNA when asked to write a draft article about Kamala Harris election campaign for Ekstra Bladet (tabloid) and Politiken (centre-left daily).*

6. **Integrating editorial values into AI systems is an ongoing challenge**
   While the above exemplifies ways of integrating editorial values into AI systems such as recommender systems and generative AI systems, they are still relatively crude and sometimes involve trade-offs between the accuracy of the systems and journalistic concerns. More elegant ways of integrating values in AI systems are, however, possible in principle, and alignment research is underway both in academia and industry that is likely to enable advances that can be exploited by news publishers, as exemplified in the box below.

ONGOING RESEARCH ON EDITORIAL ALIGNMENT OF RECOMMENDER SYSTEMS

To encourage research in novel ways of integrating editorial values into news recommender systems, PIN and JP/Politikens Media Group open sourced a very large dataset for training and evaluation of news recommenders in early 2024. The dataset was used as the basis for the ACM RecSys Challenge 2024 where 200 academics and data scientists competed in devising the best news recommendation methods. These were all evaluated for news flow effects. The challenge both showed that news recommenders ability to predict reader interests can be improved even further and confirmed PINs conclusions that recommender systems have significant and diverse news flow effects that require constraining their behavior to ensure editorial alignment and prevent audience fragmentation. While the RecSys Challenge has concluded, the dataset is still available for research purposes (see recsys.eb.dk).

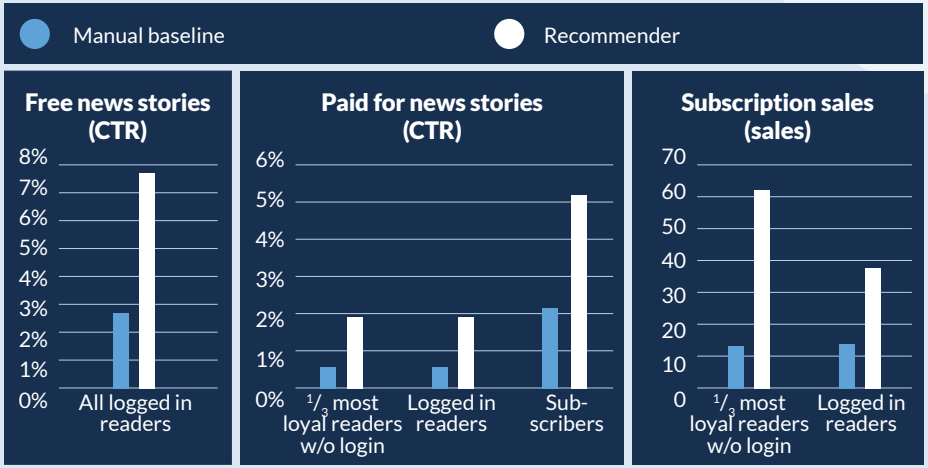
BUSINESS FINDINGS AND INSIGHTS

7. AI can increase the revenues of news publishers

AI systems can increase the revenues of news publishers in multiple different ways when used in specific applications that affect the revenue drivers in news publishers' business models. Examples include the use of AI-produced metadata to activate news content in new ways that increase engagement (e.g. in non-personalized "read more-lists" with related articles) and to generate semantic signals about readers' interests that enabled more effective targeting of advertising.

By far the most effective use case found by the PIN project, however, was the introduction of recommender systems which A/B tests showed to be directly related with significant increases in engagement (and by extension traffic and advertising inventory) and article related subscriptions sales, as illustrated in the figure below. When rolled out at scale, these effects declined, as some users declined profiling and because of constraints related to alignment with the editorial profile (see 5 above), but still increased free reading by 80%, paid reading by 35%, and subscription sales by 38% (Kruse et al 2023).

Figure: *Business effects of recommender systems compared to non-personalized methods for content selection*



Source:. *Kruse et al 2023. A/B test with duration of one week, data from one horizontal/swipable list of article suggestions on ekstrabladet.dk.*

8. Generative AI can produce tools that assist with tasks in the editorial workflow

In late 2022, generative AI emerged as a capability that was assessed ready for applied experimentation. Many actors (including PIN) have since worked on turning that capability into tools that can increase productivity and enable entirely novel approaches in news content discovery and multi-modal story telling. However, as tools build on generative AI are being introduced in newsrooms and other departments just now, neither PIN or other projects have yet been able to study productivity effects at scale (or to examine the effects that applications of generative AI that aim to enrich news coverage and storytelling might have on revenues). However, early tests indicate that generative AI currently is an assistive technology that can aid journalists in their work by providing among other things writing assistance and drafting of articles and story elements. Potentially this will make them more efficient when fully introduced in editorial workflows without enabling full automation of journalistic tasks, due to their varying alignment with editorial tone of voice and the dangers of potential hallucinations. However, further studies of rolled out tools are needed to confirm these early indications, and to assess the scale of productivity effects.

9. AI leads to dependencies on tech giants in some AI domains and not in others

AI spans multiple technologies that are all based on machine learning but characterized by very different needs for computing power and data. As news publishers begin to introduce AI, a complex web of weak and strong dependencies of publishers on tech companies appears poised to emerge.

In some AI domains (such as recommender systems), news publishers can develop and train their own models end-to-end with access to investment and competencies in data science, thus avoiding dependencies on tech companies. In other domains (such as metadata production), news publishers can create their own applications based on open-source models that are often (but not always) produced by tech giants. In yet other domains (such as generative AI), creating applications owned by publishers is increasingly infeasible, as foundation models that applications must build on are extremely expensive to develop and the best foundation models are proprietary (e.g. the GPT-models from OpenAI, Claude from Anthropic or Gemini from Google). While the emergence of new strong and week dependencies of publishers on tech companies appears poised to emerge (and did emerge around the AI systems devel-

oped on PIN), the exact structure and consequences of these dependencies are not yet clear. Part of the reason for this, is that the relationship between open-source models (such as DaBert from BotXO and Llama from META) which create weak dependencies, and proprietary AI models (such as GPT from OpenAI), which create harder dependencies is still evolving. Further, some countries are considering investing in their own foundation models in projects such as NorwAI, SweGPT and Dansk Sprogmodelskonsortium which could potentially decrease hard dependencies of news publishers on tech companies.

THE EVOLVING DEPENDENCY OF PINS AI SYSTEMS ON TECH COMPANIES

PIN had a stated goal of producing AI systems that were not dependent on the tech giants. To seek do so – and to address domain specific challenges related to news publishing – PIN initially in-sourced the design and training of all AI systems. From this point of departure, we found it possible and reasonable to train and productionize well-optimized recommender systems end to end, and to fine tune AI systems for metadata production in Danish language news publishing based on open-source foundation produced by tech giants (BERT from Google and Roberta from META) as well Danish actors (e.g. DaBert from BotXo and Ælæctra from Malte Højmark-Bertelsen) without creating strong dependencies.

In late 2021, we considered training our own foundation models to improve Danish language metadata production and enable experimentation with Danish language text generation. However, in 2022 a paradigm shift occurred in AI, related to natural language processing and generation with the introduction of the GPT3-generation of generative AI models that made the training of competitive foundations models beyond the reach of PIN (and most actors in the rest of the world). The development of editorial tools with a quality that allows for meaningful industrial application now requires use of proprietary models from companies such as OpenAI (GPT), Anthropic (Claude) and Google (Gemini), creating a new dependency on tech giants. While the emergence of open-source alternatives in 2023 and 2024 (e.g. from Meta and Mistral) can perhaps reduce dependency on any one tech giant, the use of generative AI is poised to introduce new dependencies on tech giants for news publishers – unlike the use of AI personalization and metadata production.

USER-RELATED FINDINGS AND INSIGHTS

10. Users interpret recommender systems via their experiences on social media and expectations to the publisher

When evaluating recommender systems in news, users draw on their expectations to news media in combination with their experiences of personalization on social media. Users are ambivalent about news personalization: They express concerns about potential democratic impacts including reduced exposure diversity, missing out of important news, and production of fragmented publics. At the same time, they 'feel seen' when recommenders present content matching their preferences and annoyance when it does not (Einarsson, Petrucci, Hartley, Lomborg & Kruse, in press). Despite that users recognize the positive potential of discovering relevant news more seamlessly, they are convinced that AI systems are implemented for commercial purposes, and not to improve their experiences or promote democratic ideals.

11. The (perceived) introduction of personalization changes the relationships between news publishers and readers

A central challenge of recommender systems is the reduction of editorial control and transparency of exposure. That is not per se about a black-boxed technology, but rather that when an algorithm is deployed at scale, tracing its actions can only be done on an aggregate level. Users also recognize this algorithmic mediation from other social media or e-commerce, thus also their role of shaping the model to cater to their needs. In situations where exposed recommendations are erroneous or imprecise, users react with annoyance, but they also speculate about their role in producing this output, instead of appreciating the editorial or technical reasons behind it. In other words, users take on a role as co-creators of their news flows changing the traditional division of work between curating editors and users.

12. Long-term effects are still hard to measure and understand

Findings from the controlled online experiment performed in the project (Einarsson, Lomborg & Helles, 2024) showed that recommender systems tended to expose users to more entertaining topics, more right-wing politicians, and a less diverse composition. The differences between the experimental group exposed to recommendations and the non-exposed control group were significant, but across all variables, the effects provided were small and did not transfer to significant behavioral changes aside from larger click rates. This finding attests that biases in exposure should be controlled in everyday practice, but also that changing news readers remains hard.

EFFECT AND USER STUDIES OF RECOMMENDER SYSTEMS (PIN)

PIN carried out quantitative effect studies of the news flow effects on news recommender systems implemented at scale on ekstrabladet.dk (the largest Danish news site) and user studies via workshops with users of Ekstra Bladet. Insights from these studies are the basis for insights #10-12 and fully reported in Lomborg et al (in press) and Einarsson et al (2024).

PROCESS AND ORGANIZATION FINDINGS AND INSIGHTS (ACROSS THE FOUR DOMAINS)

13. Interdisciplinary collaboration is necessary for developing healthy AI systems

AI systems are not only deeply technical, they also impact both the editorial and business aspects of news publishing, as well as the relationship between publishers and users. Designing and evaluating these systems requires ongoing collaboration across various disciplines and professions, which can be challenging due to limited resources or organizational structures that do not support interdisciplinary and cross-sectoral work. Even in well-resourced and established organizations, the challenge remains: translating complex sets of human values and judgements into quantifiable metrics that AI systems can optimize.

14. Knowledge brokers are essential for the successful adoption of AI

Successful AI system design in news publishing hinges not only on interdisciplinary collaboration but also on the presence of knowledge brokers—individuals or teams who can bridge the gap between technical experts, editorial staff, and business strategists. These knowledge brokers are crucial for translating domain knowledge and organizational goals into actionable AI design principles. However, to support this role effectively, substantial financial investment and a robust disciplinary framework are needed. This includes dedicated funding for interdisciplinary projects, long-term collaboration mechanisms, and the establishment of institutional structures that support ongoing dialogue across different fields. Without these, even the most well-intentioned efforts may struggle to align AI systems with the diverse and contingent values of news organizations.

15. Cross-industry knowledge sharing fosters responsible innovation

Knowledge sharing and cross industry collaboration can significantly aid innovation in the AI domain for news publishers. The organizations involved in PIN have experienced these effects first-hand as knowledge sharing with each other and from informal networks spawned by events organized by PIN (such as Kunstig Intelligens i Medierne and the Nordic AI in Media Summit which was organized in collaboration with the Nordic AI in Journalism Network) have been central to the AI systems and insights produced by PIN.

ORGANIZATION AND PROCESS STUDIES ON PIN

PIN carried out ethnographic studies over a period of 6 months where researchers observed work with developing AI systems for metadata production and recommender systems in Ekstra Bladet's development unit, centered on the AI innovation team. The results of these studies are the basis for insight #13-14 and more rully reported in Thylstrup and Møller-Harley (2024).

6. RECOMMENDATIONS TO NEWS PUBLISHERS

Based on learnings from the PIN project, we offer a set of recommendations for news publishers who are introducing AI into their work processes. The recommendations build on the insights presented in the previous section and aim to encourage effective and responsible use of AI among news publishers.

While the recommendations are relevant to news publishers of all sizes, not all news publishers are able to invest in designing or adapting AI systems themselves. For that reason, this section includes a separate set of additional recommendations to publishers who decide to source ready-made AI systems from external vendors.

USER-RELATED FINDINGS AND INSIGHTS

1. **Monitor developments in AI technology continuously**
   AI is a diverse and fast evolving field with the consequence that technologies age fast and both opportunities and challenges change often. To enable effective and responsible use of AI, news publishers should monitor developments in AI continuously.

WHERE TO START AND HOW TO DIVE IN DEEP?

Section 8 in this report contains a list of resources that offer a place to start or continue for both new commers and experiences practitioners in this field.

2. **Don't just copy AI solutions from companies in other sectors**
   Even though AI is often called a general-purpose technology, all AI systems are not equally suited for all domains. They are both technically and normatively adapted to the domain they are developed for. For that reason, although news publishers can draw inspiration from AI use in other media sectors, they most often cannot use the types of recommender systems or generative systems that were developed for use in sectors such as ecommerce, social media or video streaming. News publishers must adapt AI to their contexts and use cases which involves considering both the rapid pace of news and the need for publishers not to make factual errors in their news coverage. Fortunately, as PIN shows, there are many ways to do so, but it requires learning about the technology and investing in adaptation of AI systems for news.

3. **Introduce domain specific quality controls**
   AI is different than earlier digital technologies, as it is based on probability and predictions made based on training data, rather than a set of logics or rules in a stable environment. Accordingly, even the best generative AI systems can hallucinate, also when grounded in a news archive, and a recommender system's effect on the news flow can change with the news agenda. For that reason, AI systems require humans to monitor them and correct them when their behavior departs from expectations.

JOURNALISTIC RECOMMENDATIONS

4. **Monitor the effects of AI systems on your news content and news flow**
   AI systems are not neutral technologies – they behave differently depending on the choices made in the design process. Accordingly, different recommender systems will result in different news flows, while different prompts will result in different tones of voice in a news article. You must monitor the effects of your AI systems on the news content and news flow you produce – with editors in the loop.

5. **Align your AI systems with your editorial values**
   When AI systems are used by a news publisher, they must be aligned to the news brand's set of values. This requires making the values that are often deeply embedded in the gut of editors explicit so that they can be operationalized in ways that can be implemented in the design of an AI system. Value sensitive design (which PINs value compass is an example of, see section 3) can assist in this process.

6. **Always have editors in the loop – albeit at different levels of abstraction**
   With the technologies that are currently available, alignment cannot be fully automated, and checks for unwanted effects such as a misaligned news flow or a factual error in a generated summary, fact box or article draft are needed. The needs (and possibilities) for monitoring AI systems are, however, not the same across domains. While most uses of generative AI continue to require editors in the loop before the publication of its output (both to avoid errors and align tone of voice), recommender systems can be monitored by editors at an aggregate level with dashboards to identify news flow effects and take action when deemed needed to ensure editorial alignment.

BUSINESS RECOMMENDATIONS

7. **Invest in mature AI solutions to increase revenues – but also allow experimentation if you strive for competitive advantage**
   There are now types of AI technologies such as recommender systems and AI for metadata production that are industrially mature and have proven capable of creating business value. That means you can assess and invest in these systems with a traditional business case and with reasonable certainty decide when the ROI justifies investing in AI. However, AI is diverse and fast evolving – and sometimes business outcomes are hard to measure directly. Superior value creation or competitive advantage is, therefore, likely to require strategic investments in economically uncharted terrain.

8. **Invest in AI to increase productivity – but do not expect large returns in the short term**
   Generative AI is emerging as a technology that assists or sometimes even automate tasks in the workflow. However, the basic capabilities became industrially mature only a year or two ago. Right now, those capabilities are being integrated into new tools that need to be adopted by journalists. That is likely to increase productivity and while some gains can likely be realized in the short term, it will take time for large returns to emerge, as they require further development of generative tools and changes in the work routines in the editorial room.

9. **Manage your dependencies**
   The use of AI is likely to give rise to a new web of weak and strong dependencies of news publishers on tech companies. This is not necessarily bad, but history shows that publishers should be aware of the balance in transactions they engage in when building on technologies from tech companies, especially if the exchange involves access to the publisher's user or content data that can be used to compete with the publisher in other parts of the value chain. Accordingly, analysis of the value exchange and its potential effects on dependencies and competition in the information ecology value chain must be part of publishers' decisions to build technologies themselves or buy AI solutions from tech companies, including when decisions to build or buy involve weak dependencies to the creators of open-source solutions.

USER RELATED RECOMMENDATIONS

10. Understand what people value

Users interpret and evaluate news recommendations with reference to the editorial profile of the news outlet and experiences with personalization on social media. In both realms, users have become accustomed to being treated as data points that are tracked and traded; hence many have a skeptical position towards underlying intention with the introduction of AI systems. If recommenders or other types of AI systems are meant to improve users' experiences, understanding what users value and how systems can fulfill that is utmost important.

11. Recognize that users' behaviors and experiences are complementary

When developing AI systems, it is tempting to focus on news reading from a behavioral perspective that can be measured through clicks, dwell time, scroll depth, and so on. These can be viewed as the best indicators at hand to develop systems, but qualitative findings from this project nuance that behaviors are driven by several contextual factors and that relevance from the users' perspective is a rich, evolving, and complicated concept. We advocate that measures of qualitative experiences and quantitative behaviors should not compete, but rather complement each other in the building of effective and meaningful AI applications.

12. Manage how AI systems change relationships with publishers

Quantitative findings from this project demonstrate that recommender systems can introduce new biases and decrease diversity but to a relatively small extent, and the effects do not translate directly to behavior. This mirrors the fact that AI systems are rarely introduced as a replacement for the existing, but rather an incremental addition to the existing services. While the impacts of new interventions should be monitored, it is equally important to manage how the introduction of new or scaling of existing AI systems configure relationships between publishers and audiences. If all news reading were to be personalized, individual users would be attributed a greater responsibility to curate their news flow. Users have an intuitive understanding of how their interactions can shape their future experiences, but more importantly, this denotes a shift in genre where the authority of news media to set public agendas through curation is contested.

PROCESS AND ORGANIZATION RELATED RECOMMENDATIONS

13. Take a strategic and adequately funded approach

To build a supportive ecosystem for AI innovation, organizations should consider the following recommendations. It is necessary to allocate dedicated financial resources for interdisciplinary AI projects, including long-term funding models that go beyond short-term experimentation and pilot programs. It is important to develop institutional structures that support interdisciplinary research and development, including AI ethics committees that cut across the editorial, commercial and technical divides, cross-disciplinary working groups, and partnerships with academic institutions. Finally, there is a need for clear guidelines and metrics that translate organizational values into concrete goals for AI systems, ensuring that ethical considerations and editorial standards are fully integrated into the technical design process. By committing to these investments and structures, organizations can create the necessary foundation to develop AI systems that are both technically sound and aligned with broader organizational values.

14. Nourish interdisciplinary collaboration and knowledge brokers

To foster effective knowledge brokers, it is recommended to invest in training programs that equip individuals with both technical literacy and a deep understanding of editorial and business practices, enabling them to bridge disciplinary gaps. It is also recommended to formalize cross-functional teams that include knowledge brokers as central players, ensuring their involvement throughout the AI design and evaluation phases. And to support open dialogue by creating spaces where technical experts, editorial staff, and business leaders can regularly collaborate and discuss challenges and priorities in AI system development. By institutionalizing the role of knowledge brokers, organizations can support a value-based integration of AI systems that reflects the overarching strategies and obligations of news publishing.

15. Engage in collaborations with peers across the news industry

Even for the largest news publishers adopting AI involves navigating journalistic, technical, business and user-related challenges in uncharted territory. To do so effectively and safely, cross industry knowledge sharing and outright collaboration on shared challenges is essential. This applies to the challenges involved in technically adapting AI to the news domain, how to create business value with AI, how users respond to the introduction of AI and – perhaps most important – how to align AI with journalistic values.

RECOMMENDATION TO SMALL NEWS PUBLISHERS

One of PINs findings was that developing effective AI systems that are aligned with news publishers' values requires significant resources and competencies. Not all publishers will be able to – or should - mobilize the necessary level of resources and competencies necessary to design AI systems themselves. However, they will still be able to use the findings and recommendations laid out in this report as a point of departure for their dialogue with vendors and approach to using AI systems in practice. Boiled down, this will involve the following four guidelines when procuring AI solutions:

1. Make sure you and your vendor understand the specific characteristics of news publishing and that their solutions are built to exploit opportunities (such as the existence of rich metadata about news articles) and challenges (such as the fast pace of news or the need for factual correctness) associated with the news publishing domain

2. Make sure your vendor (or yourself) enables you to monitor the effects of the AI systems you introduce on your news content and news flow so that they are visible to you and can be acted upon if necessary

3. Demand adjustments to the AI systems provided by the vendor (or controls that allow you to configure and steer them yourself) if they produce results that are not aligned with your editorial values and profile

4. Establish longstanding interdisciplinary teams or working groups that have editorial, business and technical expertise when procuring and evaluating AI systems.

7. QUESTIONS ON THE HORIZON

While PIN has produced new insights in several aspects of news publishers' use of AI, many questions have been outside the scope or beyond the ability of the project to address. These include concrete questions, such as how do readers perceive the use of generative AI in news publishing? What are the long-term effects of recommender systems for news when used at scale? Will AI yield substantial long-term benefits, not just in terms of efficiency or profit, but in enhancing journalistic integrity and public trust? How should news publishers (re)organize to get the most from AI when it is introduced at scale? What are the long-term effects of news publishers' reliance on open source and/or proprietary digital solutions such as generative foundation models from tech giants?

Moreover, PIN has focused on how news publishers can use AI in effective and responsible ways to create value in ways that are aligned with their editorial missions. PIN has, however, not addressed the broader question of what a potential future information ecology driven by AI will look like, nor what the role of news publishers will be in this ecology.

Although much uncertainty exists about the timing and direction, at this point it appears increasingly likely that AI will transform the ways in which information is produced, distributed, and consumed over the coming decade. This transformation is poised to affect news publishers as well as their traditional democratic functions in fundamental ways. For that reason, it is essential for both news publishers and democratic societies in general that questions about the future of the information ecology, including how to ensure that the functions necessary for a healthy democracy continue to exist, are addressed by both academics, policy makers, news publishers and companies in other industries related to the information ecology.

In earlier waves of digitalization, news publishers have both benefitted from new ways of telling and distributing news stories and have been challenged by new dependencies on tech companies that they also compete intensely with. AI has the potential to impact publishers' core activities in news discovery, production, and distribution even more deeply. If this potential is realized in effective ways that are aligned with news publishers' values, then the transformation can greatly benefit both news publishers and the societies they are situated in. However, if AI leads to a reorganization of the information ecology that erodes the democratic functions that have traditionally been held be news publishers, then more sinister scenarios seem possible.

8. WANT TO LEARN MORE ABOUT AI IN MEDIA?

For that past three years and nine months PIN has communicated findings via presentations and academic papers and organized debates on AI in media to spur innovation and contribute to healthy norm-setting for AI in media. Most of this material is publicly available:

- All talks from three editions of the Kunstig Intelligens i Medierne (held in 2021, 2022 and 2023) are available on the event's YouTube channel: https://www.youtube.com/@kunstigintelligensimediern9589
- All talks from two editions of the Nordic AI in Media Summit (held in 2022 and 2023) which has been organized by PIN in collaboration with the Nordic AI Journalism Network are available on the event's YouTube channel: https://www.youtube.com/@NordicAIinMediaSummit-th2ik
- In the initial phases of PIN, we open sourced two NLP models for production of meta data on news articles: NERDA (https://github.com/ebanalyse/NERDA) and Senda (https://github.com/ebanalyse/senda)
- In the closing phases of PIN, we published a very large dataset for news recommendation (EB-NeRD) which was also used in ACM RecSys Challenge 2024. The dataset is available here: https://recsys.eb.dk/. A paper describing the dataset is available here: https://arxiv.org/abs/2410.03432

Learnings from PIN have been published in numerous academic papers. The following is a selection that could be particularly relevant to professionals working with AI in news publishing:

Einarsson, Á. M., Helles, R. and Lomborg, S. (2024)
Algorithmic agenda-setting: the subtle effects of news recommender systems on political agendas in the Danish 2022 general election
Information, Communication & Society, pp. 1–21.

Hartley, Jannie and Bonde Thylstrup, N. (2024)
The Algorithmic Gut Feeling – Articulating Journalistic Doxa and Emerging Epistemic Frictions in AI-Driven Data Work
Digital Journalism, pp. 1–19.

Kruse, Johannes, Kasper Lindskow, Michael Riis Andersen, and Jes Frellsen (2023).
Creating the next generation of news experience on ekstrabladet.dk with recommender systems.
In Proceedings of the 17th ACM Conference on Recommender Systems (RecSys '23). Association for Computing Machinery, New York, NY, USA, 1067–1070.

Mikkel Flyverbom (2024).
Theorizing data analysis platforms
digital refractions and reconfigurations of pasts, presents and futures

Kruse, Johannes, Kasper Lindskow, Saikishore Kalloori, Marco Polignano, Claudio Pomo, Abhishek Srivastava, Anshuk Uppal, Michael Riis Andersen, and Jes Frellsen (2024a).
RecSys Challenge 2024: Balancing Accuracy and Editorial Values in News Recommendations.
In Proceedings of the 18th ACM Conference on Recommender Systems (RecSys '24). Association for Computing Machinery, New York, NY, USA, 1195–1199.

Kruse, Johannes, Kasper Lindskow, Saikishore Kalloori, Marco Polignano, Claudio Pomo, Abhishek Srivastava, Anshuk Uppal, Michael Riis Andersen, and Jes Frellsen (2024b).
EB-NeRD a large-scale dataset for news recommendation.
In Proceedings of the Recommender Systems Challenge 2024 (RecSysChallenge '24). Association for Computing Machinery, New York, NY, USA, 1–11.

Einarsson, Á. M., Lomborg, Stine, Jannie Møller Hartley, Elisabetta Petrucci and Johannes Kruse:
I must have clicked on something – Users' Experiences and Evaluations of News Recommender Systems

PIN